Joshua Lederberg
School of Medicine
Stanford University
Stanford, California 94305

Rapid Calculation of Molecular Formulas from Mass Values

The calculation of molecular compositions consistent with a given range of mass values arises particularly in mass spectrometry. Although this can be a trivial exercise on the computer, it has been vexing to do by hand. Published tables, e.g., Beynon and Williams,[1] are bulky, and nevertheless cover a limited range of atom values. The values are also awkward to search, not having been sorted.

The following approach was designed for a desk calculator that ought to be available to any student. As it involves only a few additions and subtractions, it can—*horribilis dictu*—even be done by hand. Furthermore, it lends itself to real time implementation on small computers that lack high precision "divide" instructions in their hardware.

The basis of the calculation is the table, which is an ordered list of the mass numbers of the formulas for H from 0 to 10, N from 0 to 5, and O from 0 to 11. It contains only those compositions whose masses are an integral multiple of 12. Any number of C's may then be added as required.

The use of the table is best explained by a specific example, say $m = 259.09 \pm 0.001$.

Step 1. Since $259 \equiv 7$ modulo 12, 5 H's (5.03913) will be borrowed to give $m' = m + 5H = 264.129$. This is then divided into $m' = m_i + m_f$; $m_i = 264$ (= 22 × 12); $m_f = 0.129 \pm 0.001$.

Step 2. The table is searched for entries that correspond to m_f and whose mass does not exceed m_i. (m_i is expressed as $m_i/12$ = C-equivalent.) We find none in this cycle.

Step 3. We therefore remove 12 H's (12.0939) to give $m'' = m' - 12H = 252.035 \pm 0.001$. The table now has entries at 0.034 ($H_8N_4O_8$), 0.035 ($H_{10}NO_9$) and 0.036 ($H_6N_5O_5$). These will be completed in Step 4. 12 H's are again removed until m_f falls below − 0.0498, the bottom of the table. In our example, this occurs at the next cycle.

Step 4. The table entries are now completed as follows

			Add C's to make up m''	Adjust borrowed H's	Check mass (compare 259.0900 ± 0.0010)
34	0.034216	$H_8N_4O_8$ m_i = C_{16}	C_5	$C_5H_{15}N_4O_8$	259.089
35	0.035559	$H_{10}NO_9$ m_i = C_{14}	C_7	$C_7H_{17}NO_9$	259.090
36	0.036895	$H_6N_5O_5$ m_i = C_{14}	C_8	$C_8H_{13}N_5O_5$	259.092

Step 5. Various criteria of chemical plausibility can be used to filter the list. Since the valence rules allow H's to a maximum of 2 + 2C + N, none of these compositions is oversaturated. $C_5H_{15}N_4O_8$ however has an odd number of H's and may therefore represent a free radical.

If wider ranges of hetero atoms are contemplated, adjustments of blocks of 6 N (84.01844) and 12 O (191.9389) can be applied repetitively in a fashion similar to Step 3 so long as the adjusted mass allows.

In fact $m'' = m - 6N - 7H = 168.017 \pm 0.001$ leads to $C_6H_{11}N_8O_4$, $m = 259.090$. Further, $m - 12N - 7H = 83.999 \pm 0.001$. We read this as $m_i = 84$; $m_f = -0.001$ and find two entries in the table: −0.000826 (H_6NO_{10}) and 0.000510 ($H_2N_2O_6$), whose m_i however >84.

The table is arranged so as to illustrate its use in a fast computer program. A linear array with 138 cells, indexed as shown, has entries that never slip more than one position away from the value of the index. The composition values can therefore be accessed by direct lookup, obviating a table search. A card deck version of the table is available on request from the author.

This compilation is a greatly shortened form of some tables that were published some time ago.[2]

This work has been supported in part by the Advanced Research Projects Agency (contract SD-183), the National Aeronautics and Space Administration (grant NGR-05-020-004), and the National Institutes of Health (grant GM-00612-01).

[1] BEYNON, J. H., AND WILLIAMS A. E., "Mass and Abundance Tables for use in Mass Spectrometry," Elsevier, Amsterdam, 1963.

[2] LEDERBERG, J., "Computation of Molecular Formulas for Mass Spectrometry," Holden-Day, San Francisco, 1964.

Table of Mass Fractions for all Combinations[a] of H, N, O (H ≦ 10 N ≦ 6 O ≦ 11)

Index	$m_f \times 10^6$	H	N	O	=C	Index	$m_f \times 10^6$	H	N	O	=C	Index	$m_f \times 10^6$	H	N	O	=C
−49	−49787	0	2	11	17	0	0	0	0	0	0	31	31537	10	3	11	9
−45	−45765	0	0	9	12	1	510	2	5	6	14	32	32363	4	2	1	11
−38	−38554	0	4	10	18	2	1853	4	2	7	12	34	34216	8	4	8	16
−37	−37211	2	1	11	16	4	4532	2	3	4	9	35	35559	10	1	9	14
−34	−34532	0	2	8	13	5	5875	4	0	5	7	36	36895	6	5	5	13
−30	−30510	0	0	6	8	6	6385	6	5	11	21	38	38238	8	2	6	11
−25	−25978	2	3	10	17	7	7211	0	4	1	6	40	40917	6	3	3	8
−24	−24635	4	0	11	15	8	8554	2	1	2	4	41	42260	8	0	4	6
−23	−23299	0	4	7	14	10	10407	6	3	9	16	42	42770	10	5	10	20
−21	−21956	2	1	8	12	11	11750	8	0	10	14	43	43596	4	4	0	5
−19	−19277	0	2	5	9	13	13086	4	4	6	13	44	44939	6	1	1	3
−15	−15255	0	0	3	4	14	14429	6	1	7	11	46	46792	10	3	8	15
−14	−14745	2	5	9	16	15	15765	2	5	3	10	49	49471	8	4	5	12
−13	−13402	4	2	10	18	17	17108	4	2	4	8	50	50814	10	1	6	10
−10	−10723	2	3	7	13	18	18961	8	4	11	20	52	52150	6	5	2	9
−9	−9380	4	0	8	11	19	19787	2	3	1	5	53	53493	8	2	3	7
−8	−8044	0	4	4	10	20	21130	4	0	2	3	56	56172	6	3	0	4
−6	−6701	2	1	5	8	21	21640	6	5	8	17	57	57515	8	0	1	2
−4	−4022	0	2	2	5	22	22983	8	2	9	15	58	58025	10	5	7	16
−2	−2169	4	4	9	17	25	25662	6	3	6	12	62	62047	10	3	5	11
−1	−826	6	1	10	15	27	27005	8	0	7	10	64	64726	8	4	2	8
						28	28341	4	4	3	9	66	66069	10	1	3	6
						29	29684	6	1	4	7	68	68748	8	2	0	3
						30	31020	2	5	0	6	73	73280	10	5	4	12
												77	77302	10	3	2	7
												81	81324	10	1	0	2
												88	88535	10	5	1	8
(−0.049 to −0.0008)						(0 to 0.03)						(0.03 to 0.088)					

[a] Arranged so that the index for each entry agrees with 1000 × $m_f \pm 1.9$.

[Reprinted from Journal of Chemical Education, Vol. 49, Page 613, September, 1972.]